Kidney International, Vol. 55 (1999), pp. 1225–1233

Mutational analysis within the 3′ region of the *PKD1* gene

Celia Badenas, Roser Torra, José Luis San Millán, Liliana Lucero, Montserrat Milà, Xavier Estivill, and Alejandro Darnell

Departments of Genetics and Nephrology, Hospital Clínic, Institut d'Investigacions Biomèdiques August Pi i Sunyer, University of Barcelona; Medical and Molecular Genetics Center, Hospital Duran i Reynals, L'Hospitalet de Llobregat, Barcelona; and the Unidad de Genética Molecular, Hospital Ramón y Cajal, Madrid, Spain

Mutational analysis within the 3′ region of the *PKD1* gene.

Background. Autosomal dominant polycystic kidney disease (ADPKD) is one of the most common genetic diseases in humans, affecting 1 out of 1000 individuals. At least three different genes are involved in this disease. The search for mutations in *PKD1* is complicated because most of the transcript is encoded by a genomic region reiterated more proximally on chromosome 16, and no prevalent mutation has been reported.

Methods. We have screened DNA from exon 43 through exon 46 and intron 40 of the *PKD1* sequence by single-stranded conformational polymorphism (SSCP) analysis in 175 ADPKD patients.

Results. We have found 25 differences with respect to the reported *PKD1* DNA sequence, seven of which are mutations (Q4041X, Q4124X, IVS44-1G→C, IVS45-1G→A, 12801del28, R4275W, and Q4224P). We found different phenotypical expressions of the same mutation in the families studied. We have detected several common polymorphisms, and some of them cosegregate, suggesting a common origin of these alleles in *PKD1*.

Conclusions. The detection of only seven mutations in 175 unrelated ADPKD patients for this region of the *PKD1* analyzed suggests that mutations could be widespread throughout all of the gene and that a prevalent mutation is not expected to occur. The identified *PKD1* missense mutations may help to refine critical regions of the protein. Until a quicker and more sensitive method for the detection of mutations becomes available, linkage studies will continue to be the basis for the molecular diagnosis of ADPKD families.

Autosomal dominant polycystic kidney disease (ADPKD) is one of the most common genetic diseases in humans, affecting 1 out of 1000 individuals [1]. It is a multisystemic disease characterized by the progressive development and enlargement of multiple fluid-filled cysts in the kidney that may ultimately lead to end-stage renal disease (ESRD). Other manifestations include hepatic cysts, cerebral aneurysms, and cardiac valve abnormalities [2].

At least three different genes are involved in this disease. *PKD1* accounts for approximately 85% of ADPKD families [3] and is localized at 16p13.3 [4, 5]. *PKD2* accounts for most of the remaining cases and is localized at 4q13-23 [3, 6]. It has been proven that *PKD2* is a milder form of the disease [7–9]. A small number of families are unlinked to the *PKD1* and *PKD2* loci, suggesting the existence of at least a third locus that remains to be mapped [10–13].

The *PKD1* gene covers 52 kb of genomic DNA, contains 46 exons, and encodes a 14 kb mRNA. The predicted protein, polycystin, is a glycoprotein with 4302 amino acids. It contains 11 transmembrane domains, a large N extracellular tail that is supposed to be involved in cell–cell or cell–matrix interactions, and a cytoplasmic C termini. It contains a coiled-coil domain that interacts with PKD2. Because both proteins interact, it has been suggested that they may function through a common signaling pathway that is necessary for normal tubulogenesis [14–16].

Mutation detection in *PKD1* has been hampered by the fact that most of the *PKD1* gene is reiterated more proximally on the same chromosome and by the absence of hot spots for mutations [17–19]. To date, 44 mutations have been identified in the *PKD1* gene, and 16 of them are point mutations that create a premature stop codon. On the other hand, the studies performed by Qian et al and Brasier and Henske show that cells isolated from individual cysts exhibit a loss of the wild-type allele (17 to 24% of cysts) [20, 21]. These results suggest that the mutational mechanism in *PKD1* could be due to a second-hit mechanism in which a loss of the normal allele in a renal epithelial cell would cause cyst formation.

We studied 175 unrelated ADPKD patients for mutations in the 3′ region of *PKD1* (exons 43 through 46 and intron 40). Twenty-five changes were detected with respect to the reported *PKD1* DNA sequence, three of

Key words: ADPKD, gene mutation, SSCP analysis, inherited disease, cysts, end-stage renal disease.

Received for publication July 9, 1998
and in revised form October 15, 1998
Accepted for publication November 16, 1998

them being mutations that truncate the PKD1 protein, two altering the normal reading frame, two being missense mutations (one of them located in the coiled-coil domain), and 18 corresponding to amino acid and DNA variants.

METHODS

We screened genomic DNA from exons 43 through 46 and intron 40 of the *PKD1* sequence by SSCP analysis in 90 *PKD1* families (linkage analysis was performed previously) [8] and 85 unrelated ADPKD individuals. Intron 40 was studied when screening for a 51 bp deletion in exon 40 described by Peral et al [22].

Clinical details

PR43. He is a 45-year-old male with bilateral polycystic kidneys and moderate renal impairment (serum creatinine 4.5 mg/dl). He also has hepatic cysts and was diagnosed as hypertensive at 25 years of age. His father died at 42 years of age because of ESRD, and his sister is 42 years old and has ADPKD and mild renal failure (serum creatinine 1.8 mg/dl).

PR112. The proband is a 27-year-old male who was diagnosed as a newborn with polycystic kidney disease caused by palpation of enlarged kidneys. He has also Caroli's disease with hepatic fibrosis. Currently, the creatinine level is 5 mg/dl. The sonography scan discloses hepatomegaly with dilation of intrahepatic bile ducts preferentially in the right lobe, portal cavernomatosis, splenomegaly and enlarged cystic kidneys. The father is 54 years old and has polycystic liver and renal disease. He has been on renal replacement therapy for the last year. The mother and brother (24 years old) showed no evidence of ADPKD, either clinically or at ultrasound. Two uncles were diagnosed with ADPKD and started renal replacement therapy at the age of 60 and 58, respectively. The grandfather and the great grandfather died at 58 and 50 years of age, respectively, because of uremia.

PR114. The proband is a 37-year-old woman who entered ESRD at the age of 29. She was diagnosed with hypertension at the age of 25 and has undergone a failed renal transplantation. There is no other affected member in the family.

PR54. This is a four-generation family with several affected members. The affected grandmother died from uremia at the age of 69, whereas the affected siblings entered ESRD at the ages of 52, 56, and 62, respectively. All of them had hypertension and hepatic cysts. The grandchildren are 39, 40, and 41 years old, and all of them have hypertension, hepatic cysts, and normal renal function. The 40-year-old grandchild had a subarachnoidal hemorrhage caused by an intracranial aneurysm. There is one affected member in the younger generation who has neither hypertension nor renal failure.

PR22. This is a very large four-generation family. The ages of the different ESRD affected members are 45, 56, 56, 59, and 37 years old. Two members of the family were diagnosed with an intracranial aneurysm. Most of the affected members of the family are hypertensive and were diagnosed with this condition when they were in their 20s. All of the affected adult members of the family have hepatic cysts.

Molecular analysis

Screening for mutation Q4041X. Polymerase chain reaction (PCR) was performed using primers 3A3C1 and C2 (previously described) [5] and digested with *Pvu*II. This detects the C→T transition at nucleotide 12332 described in a previous study by Turco et al [23]. This substitution results in the nonsense mutation Q4041X in exon 44 and in an additional *Pvu*II restriction site.

Nonradioactive single-stranded conformational polymorphism (SSCP) analysis. PCR reactions were performed in a total volume of 25 μl containing 1 × Boehringer-Mannheim PCR Buffer, 5 mM of each dNTP, 10 pmol of each primer, 0.75 mM (for primers NN), and 1.5 mM (for the rest) $MgCl_2$, 1 U of Boehringer-Mannheim Taq High Fidelity Polymerase, 10% DMSO, and 500 ng of DNA. The annealing temperature was 55°C (for primers NN, HH, JJ, KK, and CB.3) and 60°C (for the rest). The primer sequences are as follows for exon 46: CB.1 (F-tctgtgggcttcagcacttt and R-agccagcagccttagcagt); CB.2 (F-tccaccccagcagcacttag and R-gtcctggttggccacac agc), and CB.3 (F-gatcttcccgtggcccat and R-gtgtccactccg actcca). For exon 44, they are as follows: CB.6 (F-cca gtggtccgtctttgg and R-aaggcgacaccagtgagg). For exon 43, the sequences are CB.7 (F-aggtgtgcctgctgctgtt and R-cca cctggtcgaagctagtg). Primers HH, JJ, KK, MM, NN, PP, and 3Δ3C1/C2 have been described previously [22].

Three microliters of PCR product were combined with loading buffer and were denatured and electrophoresed in a 12% nondenaturing acrylamide for four hours at 4°C (500 V). The different migrations were silver stained with the following protocol: 10% ethanol for 10 minutes, 1% HNO_3 for three minutes, two quick rinses with distilled water for 30 seconds, 0.2% $AgNO_3$ for 20 minutes, two quick rinses with distilled water, 29.6% Na_2CO_3/ 0.054% formaldehyde 37% for developing, and 10% acetic acid for 10 minutes.

Sequence analysis. PCR products were purified using the QIAquick spin PCR purification kit (Qiagen, Chatsworth, CA, USA) and automatically sequenced using the Dye Terminator Cycle Sequencing Ready Reaction (Perkin-Elmer Cetus) and an automatic sequencer (ABI310).

Computer analysis

The COILS2.2 program was used to study the coiled-coil domain of polycystin in mutation Q4224P [24]. This program predicts the structure on the basis of its similar-

Table 1. Mutations, amino acid and DNA variants in the *PKD1* gene

Name	Change	Localization	Heterozygosity
Mutations			
Q4041X	12332C→T	Exon 44	a
Q4124X	12581C→T	Exon 45	a
IVS44-1G→C		Intron 44	a
IVS45-1G→A		Intron 45	a
12801del28	28bp deletion	Exon 46	a
R4275W	13034C→T	Exon 46	0.01
Q4224P	12882A→C	Exon 46	a
Amino acid variants			
A3985G	12168C→G	Exon 43	a
I4044V	12341A→G	Exon 44	0.39
A4058V	12384C→T	Exon 45	0.13
V4145I	12644G→A	Exon 45	0.01
DNA variants			
IVS40+17del14		Intron 40	0.01
R3970R-C/T	12124C→T	Exon 43	0.03
A3991A-C/G	12184C→G	Exon 43	a
L4032L-G/C	12307G→C	Exon 44	a
IVS44+37ins5		Intron 44	a
IVS44+19delG		Intron 44	0.04
A4091A-A/G	12484A→G	Exon 45	0.39
L4136L-C/T	12617C→T	Exon 45	0.13
P4254P-C/T	12973C→T	Exon 46	a
	13135G→A	3′UTR	0.05
	13233A→G	3′ UTR	a
	13357T→C	3′UTR	a
	13364C→T	3′UTR	a
	13370G→A	3′UTR	a

[a] Changes detected in only one chromosome of the samples studied

ity to a database of known two-stranded coiled-coils. According to program instructions, the MTIDK matrix was used with scanning windows of 14, 21, and 28 residues, and 2.5 weighting for the "a" and "d" positions was used to reduce bias toward the hydrophilic, charge-rich sequences.

Sequence comparison of the *PKD1* region, including mutation R4275W in human, mouse, and pufferfish (*Fugu rubripes*), was performed using the BLAST, FASTA, and CLUSTAL-W software programs (Internet address: http:www2.wbj.ac.uk).

RESULTS

In the screening for mutations in 175 unrelated individuals, we detected 25 differences with respect to the reported *PKD1* DNA sequence (GDB accession numbers: L33243, L43618, and L43619). Seven of them correspond to nonsense, missense, frameshift, or splice site mutations. Four are amino acid variants. Six are nucleotide substitutions that do not result in an amino acidic change. Three correspond to changes in intron sequences, and five are located in the 3′UTR.

Mutations

We identified seven mutations in the *PKD1* gene (Table 1); two of them have been previously reported [19, 23, 25, 26], and the other five represent new mutations. All but one have proved to be private mutations, found only in one family of the 175 tested in this study. Six of them are point mutations that either create a premature stop codon, alter the splicing consensus sequence, or are missense mutations. The seventh mutation produces a larger protein.

Nonsense mutation Q4041X. During the screening of 120 unrelated ADPKD individuals for mutation Q4041X, we detected one subject (PR43) with the predicted abnormal heterozygous pattern. This mutation was confirmed by sequencing and was found in three pedigrees with different geographic origins and confirmed to arise independently [19]. Moreover, in patient PR43, mutation Q4041X was present in the same chromosome with variants L4032L-C, V4044, and A4091A-G (Fig. 1).

Nonsense mutation Q4124X. SSCP analysis with primers HH in individual PR114 showed an abnormal pattern. Sequencing revealed a C→T substitution at position 12582, changing Glutamine 4124 to a stop codon (CAG→TAG). This change creates *MaeI* and *StyI* restriction sites. This mutation has also been described by Daniells et al [26].

Frameshift mutation 12801del28. SSCP analysis with primers JJ in one *PKD1* patient (PR112) showed an abnormal pattern. Sequencing revealed a 28 bp not in-frame deletion. This deletion was also detected in the proband's son, but not in his unaffected one. The predicted abnormal protein would be 44 amino acids larger than normal polycystin [27].

Splice site mutation IVS44-1G→C. SSCP analysis using primers PP showed a heterozygous pattern in individual PR54. Sequencing revealed a G→C transition at the acceptor splicing site. This mutation abolishes a *Bbv*I restriction site. Family data for this patient were not available.

Splice site mutation IVS45-1G→A. SSCP analysis with primers MM showed an abnormal pattern in one *PKD1* patient (PR22). Sequencing revealed a G→A transition at the acceptor splicing site. Family study showed that IVS45-1G→A segregated with the disease. This mutation abolishes a *Bsr*I restriction site that was used to perform the family study.

Missense mutations Q4224P and R4275W. We detected two missense changes located in exon 46, linked to the affected chromosome, which could be considered as missense mutations. Q4224P is located in the coiled-coil domain of the PKD1 protein and abolishes a *Bsr*I restriction site. This enzyme was used to screen this change in controls. Q4224P was not found in 365 normal chromosomes (190 control chromosomes and 175 belonging to the ADPKD patients; Table 2). R4275W was found linked to the *PKD1* chromosome in two unrelated families, and in both cases, this missense change was linked to four variants (A4091A-G, V4058, V4044, and L4136L-T). Mutation R4275W abolishes a *Msp*I restric-

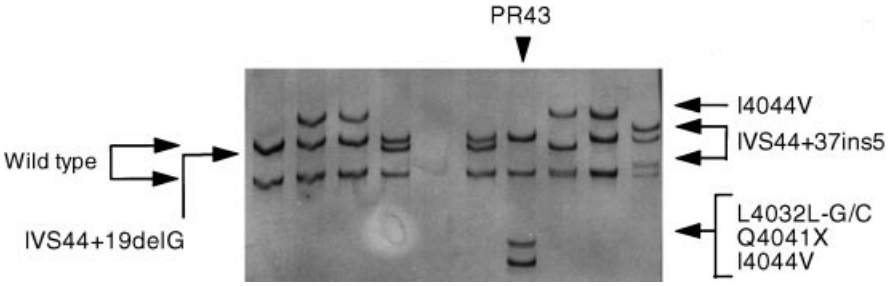


Fig. 1. Single-strand conformational polymorphism (SSCP) analysis of the CB.6 polymerase chain reaction (PCR) product corresponding to exon 44 and intron 44 of the *PKD1* gene in patients with autosomal dominant polycystic kidney disease (ADPKD). Abnormal fragments correspond to L4032L-C, Q4041X, V4044, IVS44+19delG, and IVS44+37ins5. PR43 presents the pattern of a normal individual plus two extra bands representing the mutant chromosome (carrying mutation Q4041X and the L4032L-C and IVS44+19delG variants). The fact that the normal pattern is present in this patient clearly indicates that the three detected changes are on the same chromosome.

Table 2. Rare amino acid and DNA variants in the *PKD1* gene

Change	ADPKD patients	Linked to *PKD1*	Controls	Other DNA changes
Q4224P	1/175	Y	0/95	
A3991A-C/G	1/175	Y		**12801del28**
L4032L-G/C	1/175	Y		**Q4041X**, V4044, A4091A-G
P4254P-C/T	1/175	Y	a	
IVS44ins5	1/175	Y		IVS45-1G→A
A3985G	1/175	N		
13233A→G	1/117	Y	0/71	
13370G→A	1/117	Y		
13364C→T	1/117	N		
13357T→C	1/117	N		
R4275W	2/175	Y	0.150	A4091A-G, V4058, V4044, L4136L-T
IVS40+17del14	2/95	Y/N	0/46	
V4145I	3/175	U	1/173	
R2970R-C/T	5/175	Y/N		

The number of ADPKD patients in which the changes are detected is shown and also if they are related to the affected chromosome. Abbreviations are: Y, linked to the *PKD1* allele; N, not linked; U, only found in isolated ADPKD individuals.
[a] Detected in previous study [22]

tion site, and it was used to screen this mutation in 150 controls. R4274W was not found in 459 normal chromosomes (300 control chromosomes and 159 normal chromosomes belonging to ADPKD patients).

Amino acid and DNA variants

We identified four missense changes and 14 amino acid variants with respect to the reported *PKD1* DNA sequence (Table 1). The amino acid variants were found in control chromosomes. A3985G was found only once in the population studied but was linked to the normal allele (Table 2). V4145I was found in three ADPKD individuals, and linkage analysis was not possible. This change can be detected by digestion with *Dpn*II, and it was used to screen 173 control samples, being found in one of them. A4058V and I4044V were detected with a heterozygosity of 0.13 and 0.39, respectively.

There were 14 DNA variants detected by SSCP analysis (Table 1). Three of them were located in introns: IVS40+17del14 was found in two *PKD1* families, in one linked to the pathogenic allele, and in the other linked to the wild-type *PKD1* gene (Table 2). IVS44+19delG was found in another study [18] and was observed at a frequency of 0.04. IVS44+37ins5 was only found in a *PKD1* family, linked to the affected chromosome and to mutation IVS45-1G→A (family PR22). L4032L-C was detected in a *PKD1* patient (PR43) and linked to mutation Q4041X and variants V4044 and A4091A-G (Fig. 1). R3970R-C/T was found in five unrelated patients belonging to 175 unrelated ADPKD individuals. In three of these cases, linkage analysis was possible, and this change was linked to the affected chromosome in one family, whereas in the other two, it was linked to the normal chromosome (Table 2).

Three DNA variants were relatively common: two conserve the amino acidic sequence (L4136L-C/T and A4091A-A/G), and the third one is located in the 3′UTR (13135G→A). We detected six nucleotide substitutions that were found only once in ADPKD patients and do not alter the amino acid sequence or are located in the 3′UTR: A3991A-C/G, P4254P-C/T (previously described) [22], 13370G→A, and 13233A→G were found associated with the *PKD1* chromosome, whereas 13357T→C and 13364C→T with the normal allele. Several polymorphisms detected in this study can be used in linkage analysis as intragenic markers. 13135G→A has a hetero-

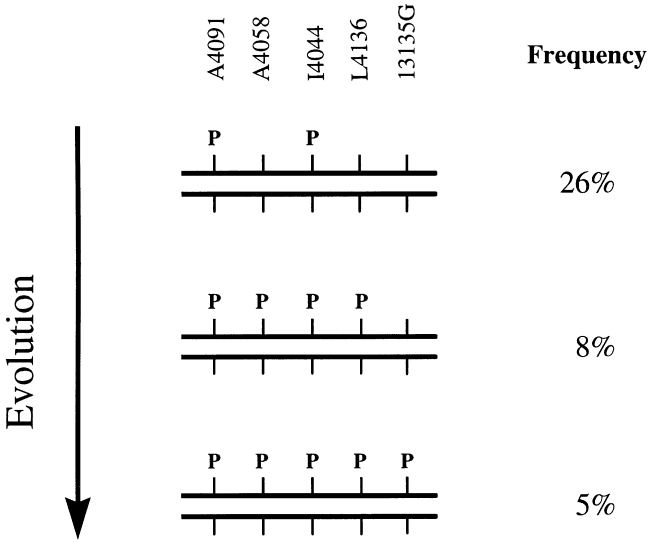


Fig. 2. Proposed evolution of *PKD1* chromosomes carrying variants A4091A-A/G, A4058V, I4044V, L4136L-C/T, and 13135GA. Frequencies for these haplotypes are shown on the right.

zygosity of 0.05, A4058V and L4136L-C/T of 0.13, and A4091A-A/G and I4044V of 0.39 in the Spanish population (Table 1).

Five of these variants seem to be inherited together: A4091A-G and V4044 are always found in the same chromosomes. Whenever V4058 and L4136L-T are detected, the previous two are also present. If 13135G→A is present in a chromosome, the other four variants are detected too. This sequence of changes led us to think that a founder chromosome exists in which there has been an accumulation of changes (Fig. 2).

DISCUSSION

Mutation detection in the *PKD1* gene is complicated because of the fact that it is a very large gene, with all but 3.5 kb reiterated more proximally on the same chromosome and a small number of mutations being found in more than one family. To date, only 44 mutations have been identified in the *PKD1* gene, most of them resulting in a truncated protein (Table 3 and Fig. 3). These results suggest that ADPKD may be caused by a loss of function mechanism and that cyst formation is the result of another mechanism inactivating the wild-type copy of *PKD1*. This mechanism could explain the phenotypical variation detected in *PKD1* individuals, in which members of the same family exhibit a wide range of phenotypical variation [27, 28].

We performed an SSCP analysis from exons 43 to 46 and intron 40 in the *PKD1* gene and detected seven mutations, four amino acid variants, and 14 DNA variants. The mutations detected in this study were mainly frameshifting or stop mutations, eliminating significant portions of the gene and, hence, clearly were pathogenic changes. In addition to these mutations, two missense mutations were also identified in exon 46 (Q4224P and R4275W). Only mutation R4275W has been found in more than one family studied here, although mutations Q4041X and Q4124X have also been reported previously [19, 23, 26]. Moreover, mutation R4275W has been found linked to A4091A-G, V4058, V4044, and L4136L-T substitutions in two unrelated families and in both cases

Table 3. Mutations described in the *PKD1* gene

Name	Location	Reference
Nucleotide substitutions (missense/nonsense)		
W1874X	Exon 15	[29]
Q2900X	Exon 23	[29]
L2993P	Exon 25	[17]
Q3016R	Exon 25	[17]
E3020X	Exon 25	[17]
L3510V	Exon 35	[17]
Q3513X	Exon 35	[17]
E3631D	Exon 36	[22]
M3677T	Exon 38	[36]
Y3818X	Exon 41	[28]
Q3837X	Exon 41	[22]
Q4010X	Exon 44	[18]
W4011X	Exon 44	[29]
R4020X	Exon 44	[33]
E4024X	Exon 44	[18]
G4031D	Exon 44	[18]
Q4041X	Exon 44	[19, 23, this study]
Q4059X	Exon 45	[26]
C4086X	Exon 45	[32]
Q4124X	Exon 45	[26, this study]
Y4126X	Exon 45	[36]
Q4224P	Exon 46	[This study]
R4227X	Exon 46	[22]
R4275W	Exon 46	[This study]
Nucleotide substitutions (splicing)		
IVS39+1G→C	Intron 39	[17]
IVS44+1G→C	Intron 44	[5]
IVS44−1G→C	Intron 44	[This study]
IVS45−1G→A	Intron 45	[This study]
Deletions		
4077delT	Exon 15	[29]
4247del2	Exon 15	[29]
5225delA	Exon 15	[29]
6785del17	Exon 15	[29]
IVS31+25del19	Intron 31	[17]
IVS39266del72	Intron 39-Exon 40	[22]
8657delC	Exon 23	[17]
9299delC	Exon 25	[17]
10262del2kb	Introns 30–34	[5]
11457del15	Exon 39	[22]
10708del5.5kb	Intron 34-Exon 46	[5]
IVS43+14del20	Intron 43	[37]
IVS43+17del18	Intron 43	[37]
12262del2	Exon 44	[18]
12739delA	Exon 46	[17]
12801del28	Exon 46	[27, this study]
Insertions		
10947insT	Exon 36	[22]
11549ins10	Exon 40	[36]
12416ins20	Exon 45	[26]
12511insG	Exon 45	[26]

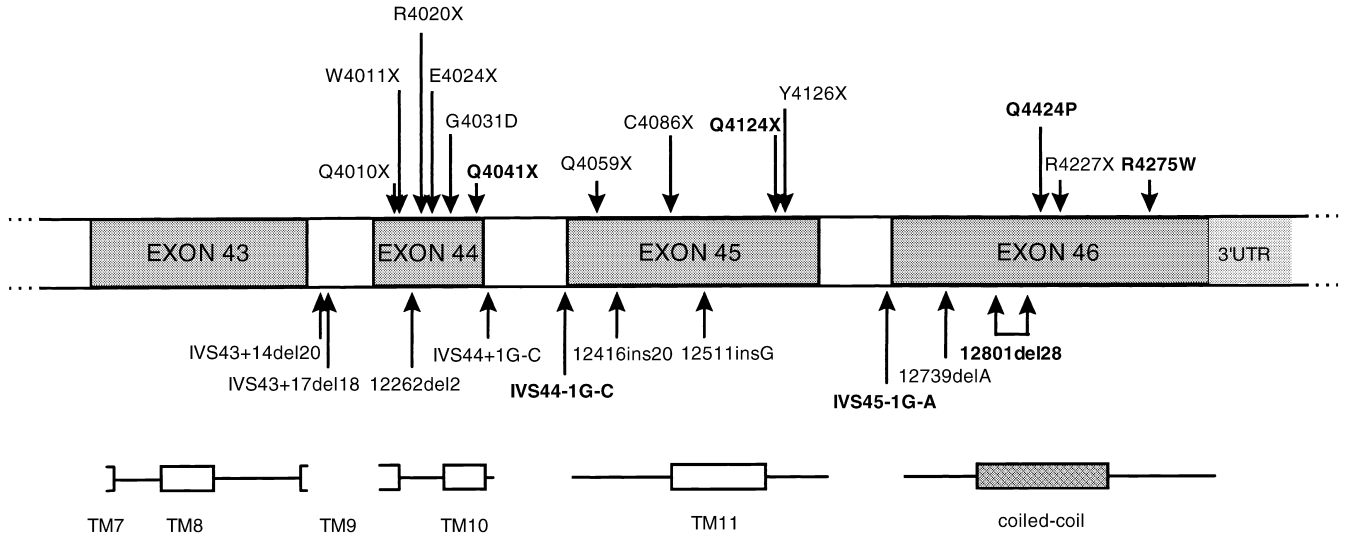


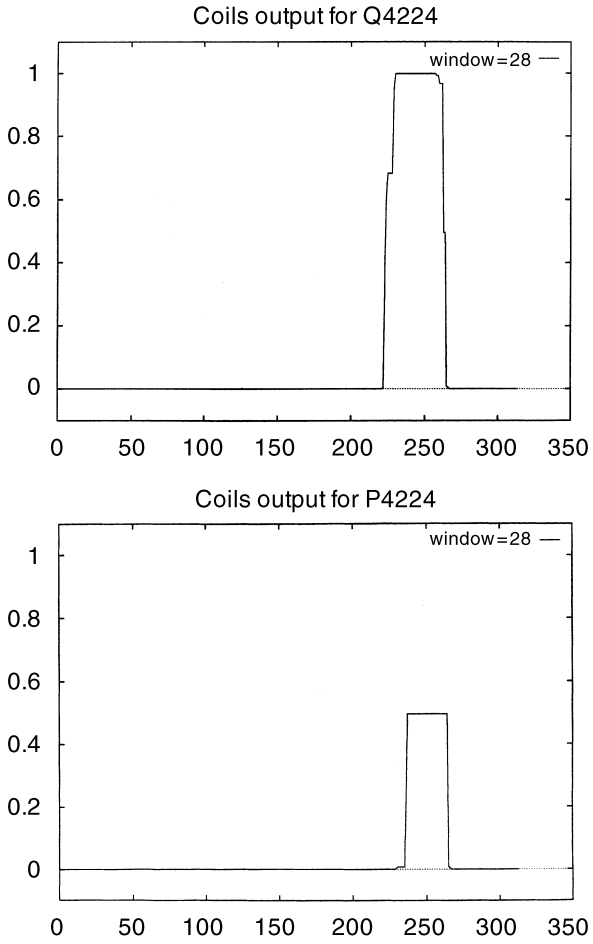
Fig. 3. Map of the region of the *PKD1* transcript analyzed (exons 43 through 46) showing the locations of described mutations and of the coiled-coil domain. Mutations described in this article are shown in bold. Exons are shown in dark gray, introns in white, and the 3′UTR in pale gray. PKD1 domains as described by Sandford et al are shown at the bottom [38]: TM, transmembrane domains (white boxes) and coiled-coil domain (filled box). Mutation 10708del5.5kb, comprising from intron 34 to exon 46, is not shown.

were linked to the affected chromosome. Because this missense change was detected in both families, a common ancestry could be postulated. In this and other studies performed thus far, no prevalent mutation has been found in the *PKD1* gene, and most of them correspond to private mutations. This implies that mutations in this gene could be found anywhere in the *PKD1* gene, as three of these studies were performed in the reiterated part of the gene, and no prevalent mutations were found there either [17, 29, 30].

Two missense mutations were identified: Q4224P and R4275W. Missense mutations are important to determine critical regions of the protein. It was difficult to determine whether these substitutions were pathogenic mutations or rare polymorphisms, especially as the entire *PKD1* gene was not screened to exclude other changes. However, in all cases in which samples were available, segregation of the mutation with the disease was demonstrated, and they were not found in a sizable screen of normal individuals. Moreover, a comparison of these residues in other species in which the *PKD1* gene has been sequenced shows that they are conserved and Q4224P is localized in a relevant structure, the coiled-coil domain of polycystin. Qian et al identified a probable coiled-coil structure within the C terminus of the *PKD1* gene product [15]. It is known that these domains in other proteins may form multimeric complexes with both coiled-coil and noncoiled-coil structures. Using the yeast two hybrid system, it was demonstrated that the coiled-coil domain of polycystin binds specifically to the C-terminus of the *PKD2* gene product, probably through another coiled-coil domain present in this protein, allowing the formation of heterodimers. The coiled-coil domain of *PKD1* comprises five consecutive heptad repeats that have the pattern characteristic of α-helical coiled-coil motifs (Fig. 4) [24]. Mutation Q4224P, identified in this work, is located in the position "d" of the second heptad. The presence of a proline residue in the heptad is likely to disrupt the α-helical structure. In fact, proline is a residue that is rarely found at any position of the known coiled-coil heptad repeats [24].

To check if Q4224P mutation affects the coiled-coil forming probability of this structure, both alleles were analyzed with the COILS program, which predicts α-helical coiled-coil domains in proteins [24]. As shown in Figure 4, the presence of the proline residue at the position "d" of the second heptad produces a dramatic reduction of more than 50% in the coiled-coil forming probability in this motif (P of wild-type allele = 1.0, P of mutant allele = 0.496) when using the most stringent criteria (window size of 28 residues). In general, it is assumed that residues with probabilities greater than 0.5 are part of the coiled-coil segment [31]. While the individual relative probability of the glutamine residue is 0.998, that of the proline residues is only 0.007. The theoretical effect of this mutation on the motif is to reduce the number of heptads identified by the COILS program from five to the last three.

Although other experimental approaches will be necessary to assess the effect of this change, this modeling study supports the hypothesis that Q4224P is a disease-causing mutation. Other authors have reported pathogenic mutations in the *PKD1* gene, because of deletions or nonsense changes that are lacking this motif [18, 22,



A LDGLSVS
B LGRLGTR
CEPEPSR
1 LQAVFEA
2 LLT**Q**FDR
3 LNQATED
4 VYQLEQQ
5 LHSLQGR

a b c d e f g

Fig. 4. COILS analysis of *PKD1* mutation Q4224P. Plots showing the probability of forming a coiled-coil structure (*y* axis, from 0 to 1) against amino acid number (*x* axis, only the carboxy termini are represented) of the wild-type allele Q4224 and the P4224. Analyses were performed as indicated in the **Methods** section with the most stringent criteria (window size 28). The predicted coiled-coil of polycystin is illustrated as heptad repeats of seven amino acids. The letters a through g designate the positions of residues within the heptad with hydrophobic residues generally found at the "a" and "d" positions. The five consecutive heptads that from the coiled-coil structure are numbered 1 through 5 (residues 4214 through 4248 of polycystin). The glutamine at position 4224 is boxed. Two additional heptads (*A* and *B*), separated from the group of five by a proline-rich block of seven amino acids (underlined), are identified when a smaller window size is used.

26, 29, 32, 33]. To our knowledge, this is the first case described of a missense mutation within the coiled-coil domain of polycystin.

The other missense mutation results in the substitution of an arginine residue for a tryptophan at codon 4275, close to the cytoplasmic C terminus of polycystin. This nonconservative change removes a positively charged residue and may be significant. This mutation is located in a region of the protein in which no homology with other proteins or motifs has been identified. However, comparison of this residue to the *PKD1* sequence from other species shows that it is highly conserved. If we look at the corresponding sequence in mouse (GDB accession number U70209) and pufferfish (GDB accession number AF013614), we can see a certain homology at that region (human amino acids 4270 to 4277, -PSRLA*R*AS-; mouse 4261 to 4268, -PSRLS*R*AS-; pufferfish 4493 to 4500, -DSHLP*R*TN-). An arginine at the same position (italicized) is conserved in the three species, indicating that substitution of this charged residue may be important.

Ultimately, a functional test for polycystin, in which missense changes can be assayed, will be required to determine whether these two changes are pathogenic mutations.

On the other hand, we have detected substitutions A4091A-G, V4044, V4058, L4136L-T, and 13135G→A that are inherited together in some individuals, either in the affected or the normal chromosomes. This suggests the existence of a founder chromosome. We propose that the sequence of changes during evolution has been the following: A4091A-G and V4044, afterward V4058 and L4136L-T, and finally 13135G→A (Fig. 2). As these changes do not always belong to the affected chromosome, they imply that mutations in the *PKD1* gene may have arisen from different mutational events, as different variants are linked to the affected chromosomes. The association of these variants with a mutation has been reported by other authors. Rossetti et al described a family in which mutation R4020X was linked to V4044 and V4058, but it is possible that they did not detect A4091A-G and L4136L-T [33]. Constantinides et al detected that V4058 and A4091A-G were linked in 12% of chromosomes, whereas A4091A-G was found in 15% of chromosomes [34]. As in our study, they suggest that

V4058 occurred more recently than the establishment of A4091A-G. On the other hand, they did not detect polymorphisms V4044, L4136L-T, or 13135G→A. Watnick et al detected also a stretch of changes in exon 23 and 25 when studying the duplicated portion of *PKD1* [30]. The authors suggest that mutations in *PKD1* may result from gene conversion between *PKD1* and its homologues [35]. This cannot be an explanation for our findings as we are studying the nonreiterated part of *PKD1.* Another example of a stretch of changes in one chromosome is found in individual PR43 (Fig. 1). This patient showed three changes in the same chromosome within a 35 bp interval: L4032L-C, Q4041X, and V4044. A4091A-G was also present in the same chromosome.

From the results presented in this study, it is obvious that the search for mutations in *PKD1* is an arduous task. A number of nucleotide changes are detected very frequently, and it is difficult to confirm if some amino acidic substitutions are mutations or only variants without functional significance. It will be interesting to know if these polymorphisms modify the disease expression. The lack of a hot spot for *PKD1* mutations makes the search for mutations difficult. From the data reported in this work and in the literature, it seems that a phenotype/genotype correlation is not evident for *PKD1.* Only when a large number of mutations are described could such a correlation be defined or could we determine whether some critical regions are linked to the different clinical features of the disease. From recently reported data it is evident that ADPKD occurs by a cellular recessive mechanism [20, 21]. It seems probable that the intrafamilial and interfamilial clinical variability of the disease may be due to a different predisposition to mutation of the normal *PKD1* allele. Another explanation for this variability may be the interaction of unknown modifier genes with PKD1.

In conclusion, we identified seven *PKD1* mutations, four amino acid variants, and 14 DNA variants in the *PKD1* gene from a total of 175 unrelated ADPKD patients. Although the screening focused on a small portion of the gene, the study suggests that a wide variability of mutations and DNA variants exists. From the mutation heterogeneity detected here and elsewhere, it seems clear that until a quicker and more sensitive method for the identification of mutations in *PKD1* becomes available, linkage analysis will continue to play an important role in the molecular diagnosis of ADPKD families.

ACKNOWLEDGMENTS

We thank the "Fondo de Investigaciones Sanitarias de la Seguridad Social" (Grant 97/2047) for partial support of this study. C. Badenas is recipient of a grant from the Instituto de Salud Carlos III (ref.98/4012). This work was presented in part at the XXXIV Congress of the European Dialysis and Transplant Association, Geneve, Switzerland, 1997.

Reprint requests to Dr. Roser Torra, S. de Nefrologia, Hospital Clínic, Villarroel 170, 08036 Barcelona, Spain.
E-mail: rtorra@medicina.ub.es

REFERENCES

1. Gabow PA: Polycystic kidney disease: Clues to pathogenesis. *Kidney Int* 40:989–996, 1991
2. Gabow PA: Autosomal dominant polycystic kidney disease: More than a renal disease. *Am J Kidney Dis* 16:403–413, 1990
3. Peters DJM, Spruit L, Saris JJ, Ravine D, Sandkuijl LA, Fossdal R, Boersma J, van Eijk R, Nrby S, Constantinou-Deltas CD, Pierides A, Brissenden JE, Frants RR, van Ommen GJB, Breuning MH: Chromosome 4 localization of a second gene for autosomal dominant polycystic kidney disease. *Nat Genet* 5:359–362, 1993
4. Reeders ST, Breuning MH, Davies KE, Nicholls RD, Jarman AP, Higgs DR, Pearson PL, Weatherall DJ: A highly polymorphic DNA marker linked to adult polycystic kidney disease on chromosome 16. *Nature* 317:542–544, 1985
5. European ADPKD Consortium: The polycystic kidney disease 1 gene encodes a 14-kb transcript and lies within a duplicated region on chromosome 16. *Cell* 77:881–894, 1994
6. Kimberling WJ, Kumar S, Gabow P, Kenyon JB, Conolly CJ, Somlo S: Autosomal dominant polycystic kidney disease: Localization of the second gene to chromosome 4q13-23. *Genomics* 18:467–472, 1993
7. Ravine D, Walker RG, Gibson RN, Forrest SM, Richards RI, Friend K, Sheffield LJ, Kincaid-Smith P, Danks DM: Phenotype and genotype heterogeneity in autosomal dominant polycystic kidney disease. *Lancet* 340:1330–1333, 1992
8. Torra R, Badenas C, Darnell A, Nicolau C, Volpini V, Revert L, Estivill X: Linkage, clinical features and prognosis of ADPKD types 1 and 2. *J Am Soc Nephrol* 7:2142–2151, 1996
9. Wright GD, Hughes AE, Larkin KA, Doherty CC, Nevin NC: Autosomal dominant polycystic kidney disease with minimal clinical expression unlinked to the *PKD1* locus. *Nephrol Dial Transplant* 8:491–494, 1993
10. Bogdanova N, Dworniczak B, Dragova DD, Todorov V, Dimitrakov D, Kalinov K, Hallmayer J, Horst J, Kalaydjieva L: Genetic heterogeneity of polycystic kidney disease in Bulgaria. *Hum Genet* 95:645–650, 1995
11. Daoust MC, Reynolds DM, Bichet DG, Somlo S: Evidence for a third genetic locus for autosomal dominant polycystic kidney disease. *Genomics* 25:733–736, 1995
12. de Almeida S, de Almeida E, Peters D, Pinto JR, Távora I, Lavinha J, Breuning M, Prata MM: Autosomal dominant polycystic kidney disease: Evidence for the existence of a third locus in a Portuguese family. *Hum Genet* 96:83–88, 1995
13. Turco AE, Clementi M, Rosetti S, Tenconi R, Pignatti PF: An Italian family with autosomal dominant polycystic kidney disease unlinked either to the *PKD1* or *PKD2* gene. *Am J Kidney Dis* 28:759–761, 1996
14. Hughes J, Ward CJ, Peral B, Aspinwall R, Clark K, San Millán JL, Gamble V, Harris PC: The polycystic kidney disease 1 (*PKD1)* gene encodes a novel protein with multiple cell recognition domains. *Nat Genet* 10:151–160, 1995
15. Qian F, Germino F, Cai Y, Zhang X, Somlo S, Germino G: PKD1 interacts with PKD2 through a probable coiled-coil domain. *Nat Genet* 16:170–183, 1997
16. Tsiokas L, Kim E, Arnould T, Sukhatme VP, Walz G: Homo- and heterodimeric interactions between the gene products of *PKD1* and *PKD2. Proc Natl Acad Sci USA* 94:6965–6970, 1997
17. Peral B, Gamble V, Strong C, Ong ACM, Sloane-Stanley J, Zerres K, Winearls CG, Harris PC: Identification of mutations in the duplicated region of the polycystic kidney disease 1 gene (*PKD1)* by a novel approach. *Am J Hum Genet* 60:1399–1410, 1997
18. Daniells C, Maheshwar M, Lazarou L, Davies F, Coles G, Ravin E: Novel and recurrent mutations in the *PKD* (polycystic kidney disease) gene. *Hum Genet* 102:216–220, 1998
19. Torra R, Badenas C, Peral B, Darnell A, Serra E, Gamble V, Turco AE, Harris PC, Estivill X: Recurrence of the *PKD1* nonsense mutation Q4041X in Spanish, Italian and British families. *Hum Mutat* 1:S117–S120, 1998

20. Qian F, Watnick TJ, Onuchic LF, Germino GG: The molecular basis of focal cyst formation in human autosomal dominant polycystic kidney disease type I. *Cell* 87:979–987, 1996
21. Brasier JL, Henske EP: Loss of the polycystic kidney disease (*PKD1)* region of chromosome 16p13 in renal cyst cells supports a loss-of-function model for cyst pathogenesis. *J Clin Invest* 99:194–199, 1997
22. Peral B, San Millán JL, Ong ACM, Gamble V, Ward CJ, Strong C, Harris PC: Screening the 3′ region of the polycystic kidney disease 1 (*PKD1)* gene reveals six novel mutations. *Am J Hum Genet* 58:86–96, 1996
23. Turco AE, Rossetti S, Bresin E, Corrà S, Gammaro L, Maschio G, Pegnatti PF: A novel nonsense mutation in the *PKD1* gene (C3817T) is associated with autosomal dominant polycystic kidney disease (ADPKD) in a large three-generation family. *Hum Mol Genet* 4:1331–1335, 1995
24. Lupas A, Van Dike M, Stock J: Predicting coiled-coils from protein sequences. *Science* 252:1162–1164, 1991
25. Badenas C, Torra R, Darnell A, Estivill X: Mutations and intragenic polymorphisms in the diagnosis of autosomal dominant polycystic kidney disease type 1. *Contrib Nephrol* 122:45–48, 1997
26. Daniells C, Maheshwar M, Lazarou L, Davies F, Coles G, Ravine D: Gene symbol: *PKD1.* Disease: Polycystic kidney disease. *Hum Genet* 102:127, 1998
27. Torra R, Badenas C, Darnell A, Brú C, Escorsell A, Estivill X: Autosomal dominant polycystic kidney disease with anticipation and Caroli's disease associated with a *PKD1* mutation. *Kidney Int* 52:33–38, 1997
28. Peral B, Ong AC, San Millan JL, Gamble V, Rees L, Harris PC: A stable, nonsense mutation associated with a case of infantile onset polycystic kidney disease 1 (*PKD1). Hum Mol Genet* 5:539–542, 1996
29. Roelfsema JH, Spruit L, Saris JJ, Chang P, Pirson Y, van Ommen GJ, Peters DJ, Breuning MH: Mutation detection in the repeated part of the *PKD1* gene. *Am J Hum Genet* 61:1044–1052, 1997
30. Watnick TJ, Piontek KB, Cordal TM, Weber H, Gandolph MA, Qian F, Lens XM, Neumann HPH, Germino GG: An unusual pattern of mutation in the duplicated portion of *PKD1* is revealed by use of a novel strategy for mutation detection. *Hum Mol Genet* 6:1473–1481, 1997
31. Lupas A: Prediction and analysis of coiled-coil structures. *Methods Enzymol* 266:513–525, 1996
32. Neophytou P, Constantinides R, Lazarou A, Pierides A, Deltas CC: Detection of a novel nonsense mutation and an intragenic polymorphism in the *PKD1* gene of a Cypriot family with autosomal dominant polycystic kidney disease. *Hum Genet* 98:437–442, 1996
33. Rossetti S, Bresin E, Restagno G, Carbonara A, Corrà S, De Prisco O, Pignatti PF, Turco AE: Autosomal dominant polycystic kidney disease (ADPKD) in an Italian family carrying a novel nonsense mutation and two missense changes in exons 44 and 45 of the *PKD1* gene. *Am J Med Genet* 65:155–159, 1996
34. Constantinides R, Xenophontos S, Neophytou P, Nomura S, Pierides A, Constantinou-Deltas C: New amino acid polymorphism, Ala/Val 4058, in exon 45 of the polycystic kidney disease 1 gene: Evolution of alleles. *Hum Genet* 99:644–647, 1997
35. Watnick TJ, Gandolph MA, Weber H, Neumann HPH, Germino GG: Gene conversion is a likely cause of mutation in *PKD1. Hum Mol Genet* 7:1239–1243, 1998
36. Turco AE, Rossetti S, Bresin E, Englisch S, Corra S, Pignatti PF: Three novel mutations of the *PKD1* gene in Italian families with autosomal dominant polycystic kidney disease. *Hum Mutat* 10:164–167, 1997
37. Peral B, Gamble V, San Millan JL, Strong C, Sloane-Stanley J, Moreno F, Harris PC: Splicing mutations of the polycystic kidney disease 1 (*PKD1)* gene induced by intronic deletion. *Hum Mol Genet* 4:569–574, 1995
38. Sandford R, Sgotto B, Aparicio S, Brenner S, Vaudin M, Wilson RK, Chissoe S, Pepin K, Bateman A, Chothia C, Hughes J, Harris P: Comparative analysis of the polycystic kidney disease (*PKD1)* gene reveals an integral membrane glycoprotein with multiple evolutionary conserved domains. *Hum Mol Genet* 6:1483–1489, 1997